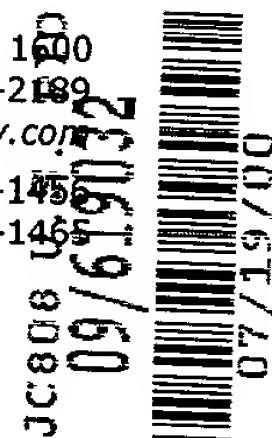




4365 Executive Drive, Suite 1000  
San Diego, CA 92121-2189  
www.graycary.com

O] 858-677-1458  
F] 858-677-1458



July 19, 2000

**Attorney Docket No.: DIVER1120-3**

Parent Serial No. 09/407,806

Parent Filing Date: September 28, 1999

**Box PATENT APPLICATION**

Assistant Commissioner for Patents

Washington, D.C. 20231

Presented for filing is a new divisional patent application of:

Applicant: DENNIS MURPHY and JOHN REID

Title: ALPHA-GALACTOSIDASE

The prior application is assigned of record to Recombinant Biocatalysis, Inc. by virtue of an assignment submitted to the Patent and Trademark Office recorded on March 8, 1996 at 7949/0419. (A certified copy of the name change from "Recombinant Biocatalysis, Inc." to "Diversa Corporation" is attached hereto.)

Enclosed are the following papers, including all those required to receive a filing date under 37 CFR § 1.53(b):

	<u>Pages</u>
Specification	25
Claim(s)	2
Abstract	1
Declaration (unsigned)	3
Revocation and New Power of Attorney (unsigned)	3
Drawing(s)	3

"EXPRESS MAIL" Mailing Label Number EL476991301US

Date of Deposit **July 19, 2000**

I hereby certify under 37 CFR 1.10 that this correspondence is being deposited with the United States Postal Service as "Express Mail Post Office to Addressee" with sufficient postage on the date indicated above and is addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231

*Mikhail Bayley*  
Mikhail Bayley

Gray Cary Ware & Freidenrich LLP

Docket No. DIVER1120-3

Box Patent Application

July 19, 2000

Page Two

Enclosures:

- Copy of Declaration (filed in the parent application, 1 page);
- Copy of Certificate of Amendment by the State of Delaware, Office of the Secretary of State, changing "Recombinant Biocatalysis, Inc." to "Diversa Corporation" (2 pages);
- Copy of Revocation and New Power of Attorney (filed in the parent application, 2 pages);
- New disclosure information, including:
  - Information disclosure statement, 2 pages
  - International Search Report (filed in the parent, 4 pages)
  - PTO Form-892 (filed in the parent, 1 page)
  - PTO-1449 (filed in the parent application, 1 page);
- Permission to Use Sequence Listing, 2 pages
  - Sequence Listing (filed in the parent application, 5 pages);
- Postcard.

This application is a divisional and claims the benefit of priority under 35 USC § 120 of U.S. application Serial No. 09/407,806, filed September 28, 1999 (pending); which is a divisional of U.S. application Serial No. 08/613,220 filed March 8, 1996 (issued on September 28, 1999 as U.S. Patent No. 5,958,751). The disclosure of the prior applications is considered part of and is incorporated by reference in the disclosure of this application.

Preliminarily, on page 1 of the specification, before line 1, insert: -- This application is a divisional and claims the benefit of priority under 35 USC § 120 of U.S. application Serial No. 09/407,806, filed September 28, 1999 (pending); which is a divisional of U.S. application Serial No. 08/613,220 filed March 8, 1996 (issued on September 28, 1999 as U.S. Patent No. 5,958,751). The disclosure of the prior applications is considered part of and is incorporated by reference in the disclosure of this application.--

This application is entitled to small entity status. Small entity status established in a previous application is still proper.

Gray Cary Ware & Freidenrich LLP

Docket No. DIVER1120-3

Box Patent Application

July 19, 2000

Page Three

Basic filing fee	\$	345.00
Total claims in excess of 20 (0 times \$9.00)	\$	0.00
Independent claims in excess of 3 (0 times \$39.00)	\$	0.00
Multiple dependent claims	\$	
Total filing fee:	\$	345.00

The payment of the filing fee is to be deferred until the Declaration is filed. Do not charge our deposit account. Please charge Deposit Account No. 50-1355 for any fees OTHER THAN the filing fee.

If this application is found to be INCOMPLETE, or if a telephone conference would otherwise be helpful, please call the undersigned at (858) 677-1456.

Kindly acknowledge receipt of this application by returning the enclosed postcard.

Please send all correspondence to:

Lisa A. Haile, Ph.D.  
Gray Cary Ware & Freidenrich LLP  
4365 Executive Drive, Suite 1600  
San Diego, CA 92121-2189  
Telephone: 858-677-1456  
Facsimile: 858-677-1465

Very truly yours,

GRAY CARY WARE & FREIDENRICH LLP



Lisa A. Haile, Ph.D.

Reg. No. 38,347

Telephone: 858-677-1456

Facsimile: 858-677-1465

006740-2306T950

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

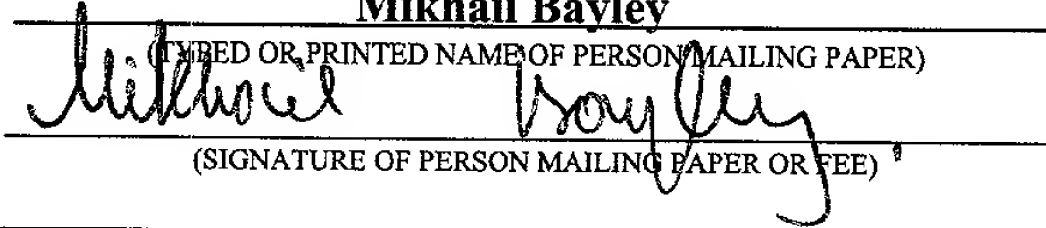
"EXPRESS MAIL LABEL NUMBER **EL476991301US**

DATE OF DEPOSIT. **July 19, 2000**

I HEREBY CERTIFY THAT THIS PAPER OR FEE IS BEING DEPOSITED WITH THE UNITED STATES POSTAL SERVICE "EXPRESS MAIL POST OFFICE TO ADDRESSEE" SERVICE UNDER 37 CFR 1.10 ON THE DATE INDICATED ABOVE AND IS ADDRESSED TO THE ASSISTANT COMMISSIONER OF PATENTS, BOX PATENT APPLICATION, WASHINGTON, D.C. 20231

**Mikhail Bayley**

(TYPED OR PRINTED NAME OF PERSON MAILING PAPER)

  
(SIGNATURE OF PERSON MAILING PAPER OR FEE)

## APPLICATION

for

UNITED STATES LETTERS PATENT

on

**ALPHA-GALACTOSIDASE**

by

**Dennis Murphy  
John Reid**

Sheets of Drawings: 3

Docket No.: DIVER1120-3

Lisa A. Haile  
Gray Cary Ware & Freidenrich LLP  
4365 Executive Drive, Suite 1600  
San Diego, California 92121-2189

09619032 071900

## ALPHA-GALACTOSIDASE

This invention relates to newly identified polynucleotides, polypeptides encoded by such polynucleotides, the use of such polynucleotides and polypeptides, as well as the production and isolation of such polynucleotides and polypeptides. More particularly, polypeptides of the present invention have been identified as glycosidases and/or  $\alpha$ -galactosidases as a result of their enzymatic activity.

In accordance with one aspect of the present invention, there are provided novel enzymes, as well as active fragments, analogs and derivatives thereof.

In accordance with another aspect of the present invention, there are provided isolated nucleic acid molecules encoding the enzymes of the present invention including mRNAs, cDNAs, genomic DNAs as well as active analogs and fragments of such enzymes.

In accordance with another aspect of the present invention there are provided isolated nucleic acid molecules encoding mature polypeptides expressed by the DNA contained in ATCC Deposit No. \_\_\_\_\_.



The following drawings are illustrative of embodiments of the invention and are not meant to limit the scope of the invention as encompassed by the claims.

Figure 1 is an illustration of the full-length DNA and corresponding deduced amino acid sequence of *Thermococcus alcaliphilus* AEDII12RA  $\alpha$ -galactosidase 18GC of the present invention. Sequencing was performed using a 378 automated DNA sequencer (Applied Biosystems, Inc.).

The term "gene" means the segment of DNA involved in producing a polypeptide chain; it includes regions preceding and following the coding region (leader and trailer) as well as intervening sequences (introns) between individual coding segments (exons).

A coding sequence is "operably linked to" another coding sequence when RNA polymerase will transcribe the two coding sequences into a single mRNA, which is then translated into a single polypeptide having amino acids derived from both coding sequences. The coding sequences need not be contiguous to one another so long as the expressed sequences ultimately process to produce the desired protein.

"Recombinant" enzymes refer to enzymes produced by recombinant DNA techniques; *i.e.*, produced from cells transformed by an exogenous DNA construct encoding the desired enzyme. "Synthetic" enzymes are those prepared by chemical synthesis.

A DNA "coding sequence of" or a "nucleotide sequence encoding" a particular enzyme, is a DNA sequence which is transcribed and translated into an enzyme when placed under the control of appropriate regulatory sequences.

In accordance with an aspect of the present invention, there are provided isolated nucleic acids (polynucleotides) which encode for the mature enzyme having the deduced amino acid sequence of Figure 1 (SEQ ID NO:4).

In accordance with another aspect of the present invention, there is provided an isolated polynucleotide encoding the enzyme of the present invention. The deposited material is a genomic clone comprising DNA encoding an enzyme of the present invention. As deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852, USA, the deposited material is assigned ATCC Deposit No. \_\_\_\_\_.

The deposit has been made under the terms of the Budapest Treaty on the International Recognition of the Deposit of Micro-organisms for Purposes of Patent Procedure. The clone will be irrevocably and without restriction or condition released to the public upon the issuance of a patent. This deposit is provided merely as convenience to those of skill in the art and is not an admission that a deposit would be required under 35 U.S.C. §112. The sequence of the polynucleotide contained in the deposited material, as well as the amino acid sequence of the polypeptide encoded thereby, are controlling in the event of any conflict with any description of sequences herein. A license may be required to make, use or sell the deposited material, and no such license is hereby granted.

The polynucleotide of this invention was originally recovered from a genomic gene library derived from *Thermococcus alcaliphilus* AEDII12RA, of the genus *Thermococcus*. AEDII12RA grows optimally at 85°C at pH 9.5.

Accordingly, the polynucleotides and enzymes encoded thereby are identified by the organism from which they were isolated, and are sometimes hereinafter referred to as "AEDII12RA- $\alpha$ -gal-18GC" (Figure 1 and SEQ ID NOS:3 and 4).

The polypeptide of the present invention shows a protein similarity of 52% and protein identity of 21% to *Dictyoglomus thermophilum* amylase.



006720-2306T950

This invention, in addition to the isolated nucleic acid molecule encoding the enzyme of the present invention, also provides substantially similar sequences. Isolated nucleic acid sequences are substantially similar if: (i) they are capable of hybridizing under conditions hereinafter described, to the polynucleotide of SEQ ID NO:3; (ii) or they encode DNA sequences which are degenerate to the polynucleotides of SEQ ID NO:3. Degenerate DNA sequences encode the amino acid sequence of SEQ ID NO:4, but have variations in the nucleotide coding sequences. As used herein, substantially similar refers to the sequences having similar identity to the sequences of the instant invention. The nucleotide sequences that are substantially the same can be identified by hybridization or by sequence comparison. Enzyme sequences that are substantially the same can be identified by one or more of the following: proteolytic digestion, gel electrophoresis and/or microsequencing.

One means for isolating the nucleic acid molecules encoding the enzyme of the present invention is to probe a gene library with a natural or artificially designed probe using art recognized procedures (see, for example: Current Protocols in Molecular Biology, Ausubel F.M. *et al.* (Eds.) Green Publishing Company Assoc. and John Wiley Interscience, New York, 1989, 1992). It will be appreciated by one skilled in the art that the polynucleotides of SEQ ID NOS:1-2, or fragments thereof (comprising at least 12 contiguous nucleotides), are particularly useful probes. Other particularly useful probes for this purpose are hybridizable fragments of the sequence of SEQ ID NO:3 (*i.e.*, comprising at least 12 contiguous nucleotides).

With respect to nucleic acid sequences which hybridize to specific nucleic acid sequence disclosed herein, hybridization may be carried out under conditions of reduced stringency, medium stringency or even stringent conditions. As an example of oligonucleotide hybridization, a polymer membrane containing immobilized denatured nucleic acids is first prehybridized for 30 minutes at 45°C in a solution consisting of 0.9 M NaCl, 50 mM NaH<sub>2</sub>PO<sub>4</sub>, pH 7.0, 5.0 mM Na<sub>2</sub>EDTA, 0.5% SDS, 10X Denhardt's, and 0.5 mg/mL polyriboadenylic acid. Approximately 2 X 10<sup>7</sup> cpm (specific activity

4-9 X 10<sup>8</sup> cpm/ug) of <sup>32</sup>P end-labeled oligonucleotide probe are then added to the solution. After 12-16 hours of incubation, the membrane is washed for 30 minutes at room temperature in 1X SET (150 mM NaCl, 20 mM Tris hydrochloride, pH 7.8, 1 mM Na<sub>2</sub>EDTA) containing 0.5% SDS, followed by a 30 minute wash in fresh 1X SET at T<sub>m</sub> -10°C for the oligo-nucleotide probe. The membrane is then exposed to autoradiographic film for detection of hybridization signals.

Stringent conditions means hybridization will occur only if there is at least 90% identity, preferably 95% identity and most preferably at least 97% identity between the sequences. See J. Sambrook *et al.*, Molecular Cloning, A Laboratory Manual (2d Ed. 1989) (Cold Spring Harbor Laboratory) which is hereby incorporated by reference in its entirety.

"Identity" as the term is used herein, refers to a polynucleotide sequence which comprises a percentage of the same bases as a reference polynucleotide (SEQ ID NO:3). For example; a polynucleotide which is at least 90% identical to a reference polynucleotide, has polynucleotide bases which are identical in 90% of the bases which make up the reference polynucleotide and may have different bases in 10% of the bases which comprise that polynucleotide sequence. As used herein, a first DNA (RNA) sequence is at least 70% and preferably at least 80% identical to another DNA (RNA) sequence if there is at least 70% and preferably at least a 80% or 90% identity, respectively, between the bases of the first sequence and the bases of the another sequence, when properly aligned with each other, for example when aligned by BLAST.

The present invention relates to polynucleotides which differ from the reference polynucleotide such that the changes are silent changes, for example the change do not alter the amino acid sequence encoded by the polynucleotide. The present invention also relates to nucleotide changes which result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference

polynucleotide. In a preferred aspect of the invention these polypeptides retain the same biological action as the polypeptide encoded by the reference polynucleotide.

The polynucleotide of this invention was recovered from a genomic gene library from *Thermococcus alcaliphilus* AEDII12RA. A gene library was generated and excisions were performed according to the protocols/methods hereinafter described.

The polynucleotide of the present invention may be in the form of RNA or DNA, which DNA includes cDNA, genomic DNA, and synthetic DNA. The DNA may be double-stranded or single-stranded, and if single stranded may be the coding strand or non-coding (anti-sense) strand. The coding sequence which encodes the mature enzyme may be identical to the coding sequence shown in Figure 1 (SEQ ID NO:3) or may be a different coding sequence, which coding sequence, as a result of the redundancy or degeneracy of the genetic code, encodes the same mature enzyme as the DNA of Figure 1 (SEQ ID NO:3).

The polynucleotides which encode for the mature enzyme of Figure 1 (SEQ ID NO:4) may include, but are not limited to: only the coding sequence for the mature enzyme; the coding sequence for the mature enzyme and additional coding sequence such as a leader sequence or a proprotein sequence; the coding sequence for the mature enzyme (and optionally additional coding sequence) and non-coding sequence, such as introns or non-coding sequence 5' and/or 3' of the coding sequence for the mature enzyme.

Thus, the term "polynucleotide encoding an enzyme (protein)" encompasses a polynucleotide which includes only coding sequence for the enzyme as well as a polynucleotide which includes additional coding and/or non-coding sequence(s).

The present invention further relates to variants of the herein described polynucleotide which code for fragments, analogs and derivatives of the enzyme having

the deduced amino acid sequence of Figure 1 (SEQ ID NO:4). The variant of the polynucleotide may be a naturally occurring or a non-naturally occurring variant of the polynucleotide.

Thus, the present invention includes polynucleotides encoding the same mature enzyme as shown in Figure 1 (SEQ ID NO:4) as well as variants of such polynucleotides which variants code for a fragment, derivative or analog of the enzyme of Figure 1 (SEQ ID NO:4). Such nucleotide variants include deletion variants, substitution variants and addition or insertion variants.

As indicated herein, the polynucleotide may have a coding sequence which is a naturally occurring allelic variant of the coding sequence shown in Figure 1 (SEQ ID NO:3). As known in the art, an allelic variant is an alternate form of a polynucleotide sequence which may have a substitution, deletion or addition of one or more nucleotides, which does not substantially alter the function of the encoded protein.

Fragments of the full length gene of the present invention may be used as hybridization probes for a cDNA or a genomic library to isolate the full length DNA and to isolate other DNAs which have a high sequence similarity to the gene or similar biological activity. Probes of this type preferably have at least 10, preferably at least 15, and even more preferably at least 30 bases and may contain, for example, at least 50 or more bases. The probe may also be used to identify a DNA clone corresponding to a full length transcript and a genomic clone or clones that contain the complete gene including regulatory and promotor regions, exons and introns. An example of a screen comprises isolating the coding region of the gene by using the known DNA sequence to synthesize an oligonucleotide probe. Labeled oligonucleotides having a sequence complementary to that of the gene of the present invention are used to screen a library of genomic DNA to identify members of the library to which the probe hybridizes.

It is also appreciated that such probes can be and are preferably labeled with an analytically detectable reagent to facilitate identification of the probe. Useful reagents include but are not limited to radioactive isotopes, fluorescent dyes or enzymes capable of catalyzing the formation of a detectable product. The probes are thus useful to isolate complementary copies of DNA from other sources or to screen such sources for related sequences.

The present invention further relates to polynucleotides which hybridize to the hereinabove-described sequence if there is at least 70%, preferably at least 90%, and more preferably at least 95% identity between the sequences. The present invention particularly relates to polynucleotides which hybridize under stringent conditions to the above-described polynucleotide. As herein used, the term "stringent conditions" means hybridization will occur only if there is at least 95% and preferably at least 97% identity between the sequences. The polynucleotides which hybridize to the hereinabove described polynucleotide in a preferred embodiment encode enzymes which either retain substantially the same biological function or activity as the mature enzyme encoded by the DNA of Figure 1 (SEQ ID NO:3).

Alternatively, the polynucleotide may have at least 15 bases, preferably at least 30 bases, and more preferably at least 50 bases which hybridize to any part of a polynucleotide of the present invention and which has an identity thereto, as hereinabove described, and which may or may not retain activity. For example, such polynucleotides may be employed as probes for the polynucleotide of SEQ ID NO:3, for example, for recovery of the polynucleotide or as a diagnostic probe or as a PCR primer.

Thus, the present invention is directed to polynucleotides having at least a 70% identity, preferably at least 90% identity and more preferably at least a 95% identity to a polynucleotide which encodes the enzyme of SEQ ID NO:4 as well as fragments thereof, which fragments have at least 15 bases, preferably at least 30 bases and most

preferably at least 50 bases, which fragments are at least 90% identical, preferably at least 95% identical and most preferably at least 97% identical under stringent conditions to any portion of a polynucleotide of the present invention.

The present invention further relates to an enzyme which has the deduced amino acid sequence of Figure 1 (SEQ ID NO:4) as well as fragments, analogs and derivatives thereof.

The terms "fragment," "derivative" and "analog" when referring to the enzyme of Figure 1 (SEQ ID NO:4) mean enzymes which retain essentially the same biological function or activity as the enzyme of SEQ ID NO:4. Thus, an analog includes a proprotein which can be activated by cleavage of the proprotein portion to produce an active mature enzyme.

The enzyme of the present invention may be a recombinant enzyme, a natural enzyme or a synthetic enzyme, preferably a recombinant enzyme.

The fragment, derivative or analog of the enzyme of Figure 1 (SEQ ID NO:4) may be (i) one in which one or more of the amino acid residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the mature enzyme is fused with another compound, such as a compound to increase the half-life of the enzyme (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature enzyme, such as a leader or secretory sequence or a sequence which is employed for purification of the mature enzyme or a proprotein sequence. Such fragments, derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.





Val, Leu and Ile; interchange of the hydroxyl residues Ser and Thr, exchange of the acidic residues Asp and Glu, substitution between the amide residues Asn and Gln, exchange of the basic residues Lys and Arg and replacements among the aromatic residues Phe, Tyr. Most highly preferred are variants which retain the same biological function and activity as the reference polypeptide from which it varies.

Fragments or portions of the enzymes of the present invention may be employed for producing the corresponding full-length enzyme by peptide synthesis; therefore, the fragments may be employed as intermediates for producing the full-length enzymes. Fragments or portions of the polynucleotides of the present invention may be used to synthesize full-length polynucleotides of the present invention.

The present invention also relates to vectors which include polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of enzymes of the invention by recombinant techniques.

Host cells are genetically engineered (transduced or transformed or transfected) with the vectors of this invention which may be, for example, a cloning vector or an expression vector. The vector may be, for example, in the form of a plasmid, a viral particle, a phage, *etc.* The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the genes of the present invention. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

The polynucleotides of the present invention may be employed for producing enzymes by recombinant techniques. Thus, for example, the polynucleotide may be included in any one of a variety of expression vectors for expressing an enzyme. Such vectors include chromosomal, nonchromosomal and synthetic DNA sequences, *e.g.*, derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids;



vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in the host.

The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art. Such procedures and others are deemed to be within the scope of those skilled in the art.

The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. As representative examples of such promoters, there may be mentioned: LTR or SV40 promoter, the *E. coli*. *lac* or *trp*, the phage lambda P<sub>L</sub> promoter and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses. The expression vector also contains a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression.

In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in *E. coli*.

The vector containing the appropriate DNA sequence as hereinabove described, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the protein.

As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as *E. coli*, *Streptomyces*, *Bacillus subtilis*; fungal cells, such as yeast; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9; animal cells such as CHO,

COS or Bowes melanoma; adenoviruses; plant cells, *etc.* The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

More particularly, the present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example; Bacterial: pQE70, pQE60, pQE-9 (Qiagen), pBluescript II KS (Stratagene), ptrc99a, pKK223-3, pDR540, pRIT2T (Pharmacia); Eukaryotic: pXT1, pSG5 (Stratagene) pSVK3, pBPV, pMSG, pSVL, SV40 (Pharmacia). However, any other plasmid or vector may be used as long as they are replicable and viable in the host.

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda P<sub>R</sub>, P<sub>L</sub> and trp. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

In a further embodiment, the present invention relates to host cells containing the above-described constructs. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection,

or electroporation (Davis, L., Dibner, M., Battey, I., *Basic Methods in Molecular Biology*, (1986)).

The constructs in host cells can be used in a conventional manner to produce the gene product encoded by the recombinant sequence. Alternatively, the enzyme of the invention can be synthetically produced by conventional peptide synthesizers.

Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual, Second Edition*, Cold Spring Harbor, N.Y., (1989), the disclosure of which is hereby incorporated by reference.

Transcription of the DNA encoding the enzymes of the present invention by higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act on a promoter to increase its transcription. Examples include the SV40 enhancer on the late side of the replication origin bp 100 to 270, a cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, *e.g.*, the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK),  $\alpha$ -factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase

with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated enzyme. Optionally, the heterologous sequence can encode a fusion enzyme including an N-terminal identification peptide imparting desired characteristics, *e.g.*, stabilization or simplified purification of expressed recombinant product.

Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but nonlimiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and pGEM1 (Promega Biotec, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced by appropriate means (*e.g.*, temperature shift or chemical induction) and cells are cultured for an additional period.

Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, such methods are well known to those skilled in the art.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, *Cell*, 23:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

The enzyme can be recovered and purified from recombinant cell cultures by methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps.

The enzyme of the present invention may be a naturally purified product, or a product of chemical synthetic procedures, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by bacterial, yeast, higher plant, insect and mammalian cells in culture). Depending upon the host employed in a recombinant

production procedure, the enzymes of the present invention may be glycosylated or may be non-glycosylated. Enzymes of the invention may or may not also include an initial methionine amino acid residue.

The polynucleotide of this invention was recovered from a genomic gene library from *Thermococcus alcaliphilus* AEDII12RA. The gene library was generated in the  $\lambda$ ZAP2 cloning vector (Stratagene). Mass excisions were performed on these libraries to generate libraries in the pBluescript phagemid. Libraries were generated and excisions were performed according to the protocols/methods hereinafter described.

### Example 1

#### Production of the Expression Gene Bank

Colonies containing pBluescript plasmids with random inserts from the organism *Thermococcus alcaliphilus* AEDII12RA were obtained from an original  $\lambda$ ZAP2 genomic library generated according to the manufacturer's (Stratagene) protocol. The clones were then excised from  $\lambda$ ZAP2 to pBluescript. The clones were excised to pBluescript according to the method of Hay and Short. (Hay, B. and Short, J. *Strategies*, 1992, 5:16.) The resulting colonies were picked with sterile toothpicks and used to singly inoculate each of the wells of 96-well microtiter plates. The wells contained 250  $\mu$ L of LB media with 100  $\mu$ g/ml methicillin, and 10% v/v glycerol (LB Amp/Meth, glycerol). The cells were grown overnight at 37°C without shaking. This constituted generation of the "Source GeneBank"; each well of the Source GeneBank thus contained a stock culture of *E. coli* cells, each of which contained a pBluescript plasmid with a unique DNA insert.

### Example 2

#### Screening for Glycosidase Activity

The plates of the Source GeneBank were used to multiply inoculate a single plate (the "Condensed Plate") containing in each well 200  $\mu$ L of LB Amp/Meth, glycerol.



This step was performed using the High Density Replicating Tool (HDRT) of the Beckman Biomek with a 1% bleach, water, isopropanol, air-dry sterilization cycle in between each inoculation. Each well of the Condensed Plate thus contained 10 to 12 different pBluescript clones from each of the source library plates. The Condensed Plate was grown for 16h at 37°C and then used to inoculate two white 96-well Polyfiltronics microtiter daughter plates containing in each well 250  $\mu$ L of LB Amp/Meth (without glycerol). The original condensed plate was put in storage -80°C. The two condensed daughter plates were incubated at 37°C for 18 h.

A '600  $\mu$ M substrate stock solution' was prepared as follows: 25 mg of each of four compounds was dissolved in the appropriate volume of DMSO to yield a 25.2 mM solution. The compounds used were 4-methylumbelliferyl  $\beta$ -D-xyloside, 4-methylumbelliferyl  $\alpha$ -D-galactoside, 4-methylumbelliferyl  $\alpha$ -D-mannopyranoside, and 4-methylumbelliferyl  $\beta$ -D-mannopyranoside. Two hundred fifty microliters of each DMSO solution was added to ca. 9 mL of 50 mM, pH 7.5 Hepes buffer. The volume was taken to 10.5 mL with the above Hepes buffer to yield a clear solution. All four umbelliferones were obtained from Sigma Chemical Co.

Fifty  $\mu$ L of the '600  $\mu$ M stock solution' was added to each of the wells of a white condensed plate using the Biomek to yield a final concentration of substrate of  $\sim$ 100  $\mu$ M. The fluorescence values were recorded (excitation = 326 nm, emission = 450 nm) on a plate reading fluorometer immediately after addition of the substrate. The plate was incubated at 70°C for 60 min and the fluorescence values were recorded again. The initial and final fluorescence values were subtracted to determine if an active clone was present by an increase in fluorescence over the majority of the other wells.

### Example 3

#### Isolation of Active Clone and Substrate Specificity Determination

In order to isolate the individual clone which carried the activity, the Source GeneBank plates were thawed and the individual wells used to singly inoculate a new plate containing LB/Amp/Meth. As above the plate was incubated at 37°C to grow the cells, the 50  $\mu$ L of 600  $\mu$ M substrate stock solution added using the Biomek. Once the active well from the source plate was identified, the cells from the source plate were streaked on agar with LB/Amp/Meth and grown overnight at to 37°C to obtain single colonies. Eight single colonies were picked with a sterile toothpick and used to singly inoculate the wells of a 96-well microtiter plate. The wells contained 250  $\mu$ L of LB/Amp/Meth. The cells were grown overnight at 37°C without shaking. A 200  $\mu$ L aliquot was removed from each well and assayed with the substrates as above. The most active clone was identified and the remaining 50  $\mu$ L of culture was used to streak an agar plate with LB/Amp/Meth. Eight single colonies were picked, grown and assayed as above. The most active clone was used to inoculate 3mL cultures of LB/Amp/Meth, which were grown overnight. The plasmid DNA was isolated from the cultures and utilized for sequencing. Colonies from this final streak onto the agar plate were also used to inoculate wells containing 250  $\mu$ L of LB/Amp/Meth. In addition, colonies containing plasmids with no inserts were used as negative controls. A 600  $\mu$ M solution of each individual substrate was made up for the purpose of determining the substrate specificity of the enzyme. Fifty  $\mu$ L of each of the four substrates were added individually to the test and control wells and assayed for activity as above. Only the wells which contained the 4-methylumbelliferyl  $\alpha$ -D-galactoside showed an increase in fluorescence indicating activity.

Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, within the scope of the appended claims, the invention may be practiced otherwise than as particularly described.



# SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: MURPHY, Dennis  
REID, John
- (ii) TITLE OF INVENTION: Alpha Glycosidase
- (iii) NUMBER OF SEQUENCES: 4
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
CECCHI, STEWART & OLSTEIN
  - (B) STREET: 6 BECKER FARM ROAD
  - (C) CITY: ROSELAND
  - (D) STATE: NEW JERSEY
  - (E) COUNTRY: USA
  - (F) ZIP: 07068
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: 3.5 INCH DISKETTE
  - (B) COMPUTER: IBM PS/2
  - (C) OPERATING SYSTEM: MS-DOS
  - (D) SOFTWARE: ASCII
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: Unassigned
  - (B) FILING DATE: Concurrently
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: HERRON, CHARLES J.
  - (B) REGISTRATION NUMBER: 28,019
  - (C) REFERENCE/DOCKET NUMBER: 331400-40
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 201-994-1700
  - (B) TELEFAX: 201-994-1744

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS
  - (A) LENGTH: 52 NUCLEOTIDES
  - (B) TYPE: NUCLEIC ACID
  - (C) STRANDEDNESS: SINGLE
  - (D) TOPOLOGY: LINEAR
- (ii) MOLECULE TYPE: cDNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGAGAGCG CTCGTCTTTC AC

52

## (2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS  
 (A) LENGTH: 31 NUCLEOTIDES  
 (B) TYPE: NUCLEIC ACID  
 (C) STRANDEDNESS: SINGLE  
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CGGAAGATCT AGGTTCCCCA TTTTCACCCC T

31

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS  
 (A) LENGTH: 1,095 NUCLEOTIDES  
 (B) TYPE: NUCLEIC ACID  
 (C) STRANDEDNESS: SINGLE  
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TTG AGA GCG CTC GTC TTT CAC GGC AAC CTC CAG TAT GCC GAA ATC CCA  
 Leu Arg Ala Leu Val Phe His Gly Asn Leu Gln Tyr Ala Glu Ile Pro  
 5 10 15

48

AAG AGC GAA ATC CCA AAG GTC ATA GAG AAG GCA TAC ATC CCA GTC ATC  
 Lys Ser Glu Ile Pro Lys Val Ile Glu Lys Ala Tyr Ile Pro Val Ile  
 20 25 30

96

GAG ACA CTG ATT AAA GAA GAA ATT CCT TTT GGG CTC AAC ATA ACG GGC  
 Glu Thr Leu Ile Lys Glu Glu Ile Pro Phe Gly Leu Asn Ile Thr Gly  
 35 40 45

144

TAT ACC TTA AAG TTC CTC CCG AAG GAT ATT ATA GAC CTC GTT AAA GGG  
 Tyr Thr Leu Lys Phe Leu Pro Lys Asp Ile Ile Asp Leu Val Lys Gly  
 50 55 60

192

GGC ATC GCG AGT GAC CTG ATA GAG ATA ATC GGA ACG AGC TAC ACG CAC  
 Gly Ile Ala Ser Asp Leu Ile Glu Ile Ile Gly Thr Ser Tyr Thr His  
 65 70 75 80

240

GCA ATA CTC CCC CTC CTC CCG CTT AGC AGA GTA GAA GCA CAA GTT CAG  
 Ala Ile Leu Pro Leu Leu Pro Leu Ser Arg Val Glu Ala Gln Val Gln  
 85 90 95

288

AGA GAT AGG GAA GTT AAG GAA GAG CTC TTC GAG CTT TCT CCA AAG GGA  
 Arg Asp Arg Glu Val Lys Glu Glu Leu Phe Glu Val Ser Pro Lys Gly  
 100 105 110

336

TTC	TGG	CTG	CCA	GAG	CTC	GCC	TAT	GAC	CCG	ATA	ATC	CCT	GCC	ATA	CTG	384
Phe	Trp	Leu	Pro	Glu	Leu	Ala	Tyr	Asp	Pro	Ile	Ile	Pro	Ala	Ile	Leu	
		115					120					125				
AAG	GAC	AAC	GGT	TAT	GAG	TAT	CTA	TTC	GCC	GAC	GGG	GAG	GCG	ATG	CTT	432
Lys	Asp	Asn	Gly	Tyr	Glu	Tyr	Leu	Phe	Ala	Asp	Gly	Glu	Ala	Met	Leu	
	130					135					140					
TTC	TCA	GCT	CAT	CTC	AAC	TCG	GCG	ATA	AAG	CCA	ATT	AAA	CCG	CTC	TAT	480
Phe	Ser	Ala	His	Leu	Asn	Ser	Ala	Ile	Lys	Pro	Ile	Lys	Pro	Leu	Tyr	
145					150					155					160	
CCA	CAC	CTT	ATA	AAG	GCC	CAA	AGG	GAA	AAG	CGC	TTT	AGG	TAC	ATC	AGC	528
Pro	His	3Leu	Ile	Lys	Ala	Gln	Arg	Glu	Lys	Arg	Phe	Arg	Tyr	Ile	Ser	
				165					170					175		
TAT	CTC	CTT	GGT	CTC	AGG	GAG	CTT	AGG	AAG	GCG	ATA	AAG	CTC	GTT	TTT	576
Tyr	Leu	Leu	Gly	Leu	Arg	Glu	Leu	Arg	Lys	Ala	Ile	Lys	Leu	Val	Phe	
			180					185					190			
GAA	GGT	AAG	GTA	ACG	CTA	AAG	GCA	GTC	AAA	GAC	ATC	GAA	GCC	GTA	CCC	624
Glu	Gly	Lys	Val	Thr	Leu	Lys	Ala	Val	Lys	Asp	Ile	Glu	Ala	Val	Pro	
		195					200					205				
GTT	TGG	GTG	GCC	GTG	AAC	ACG	GCT	GTA	ATG	CTC	GGC	ATC	GGA	AGG	CTT	672
Val	Trp	Val	Ala	Val	Asn	Thr	Ala	Val	Met	Leu	Gly	Ile	Gly	Arg	Leu	
	210					215					220					
CCT	CTT	ATG	AAT	CCT	AAG	AAA	GTG	GCG	AGC	TGG	ATA	GAG	GAC	AAG	GAC	720
Pro	Leu	Met	Asn	Pro	Lys	Lys	Val	Ala	Ser	Trp	Ile	Glu	Asp	Lys	Asp	
225					230					235					240	
AAC	ATT	CTT	CTA	TAC	GGC	ACC	GAT	ATA	GAG	TTC	ATT	GGC	TAT	AGG	GAC	768
Asn	Ile	Leu	Leu	Tyr	Gly	Thr	Asp	Ile	Glu	Phe	Ile	Gly	Tyr	Arg	Asp	
				245					250					255		
ATT	GCA	GGC	TAC	AGA	ATG	AGT	GTT	GAG	GGA	TTA	TTA	GAG	GTT	ATA	GAC	816
Ile	Ala	Gly	Tyr	Arg	Met	Ser	Val	Glu	Gly	Leu	Leu	Glu	Val	Ile	Asp	
			260					265					270			
GAG	CTC	AAC	TCG	GAA	CTG	TGC	CTT	CCC	TCA	GAG	CTG	AAG	CAC	AGT	GGA	864
Glu	Leu	Asn	Ser	Glu	Leu	Cys	Leu	Pro	Ser	Glu	Leu	Lys	His	Ser	Gly	
		275					280					285				
AGG	GAG	CTC	TAC	TTA	CGG	ACT	TCG	AGT	TGG	GCA	CCA	GAT	AAG	AGC	TTG	912
Arg	Glu	Leu	Tyr	Leu	Arg	Thr	Ser	Ser	Trp	Ala	Pro	Asp	Lys	Ser	Leu	
	290					295					300					

AGG	ATA	TGG	AGA	GAG		GAA	GGG	AAC	GCA	AGA	CTT	AAT	ATC	CTG	TCC	960
Arg	Ile	Trp	Arg	Glu		Glu	Gly	Asn	Ala	Arg	Leu	Asn	M	Leu	Ser	
305					310					315					320	
TAC	AAT	ATG	AGG	GGC	GAA	CTC	GCC	CTT	TTA	GCC	GAG	AAC	AGC	GAT	GCA	1008
Tyr	Asn	Met	Arg	Gly	Glu	Leu	Ala	Phe	Leu	Ala	Glu	Asn	Ser	Asp	Ala	
				325					330					335		
AGG	GGA	TGG	GAG	CCC	CTC	CCT	GAG	AGG	AGG	CTG	GAT	GCC	TTC	CGG	GCG	1047
Arg	Gly	Trp	Glu	Pro	Leu	Pro	Glu	Arg	Arg	Leu	Asp	Ala	Phe	Arg	Ala	
			340					345					350			
ATA	TAT	AAC	GAT	TGG	AGG	GGT	GAA	AAT	GGG	GAA	CCT	TAG				1086
Ile	Tyr	Asn	Asp	Trp	Arg	Gly	Glu	Asn	Gly	Glu	Pro	End				
		355					360					365				

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 364 AMINO ACIDS  
 (B) TYPE: POLYPEPTIDE  
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Leu	Arg	Ala	Leu	Val	Phe	Mis	Gly	Asn	Leu	Gln	Tyr	Ala	Glu	Ile	
			5						10					15	
Phe	Lys	Ser	Glu	Ile	Pro	Lys	Val	Ile	Glu	Lys	Ala	Tyr	Ile	Pro	
			20						25					30	
Val	Ile	Glu	Thr	Leu	Ile	Lys	Glu	Gln	Ile	Pro	Phe	Gly	Leu	Asn	
			35						40					45	
Ile	Thr	Gly	Tyr	Thr	Leu	Lys	Phe	Leu	Pro	Lys	Asn	Ile	Ile	Asp	
			50						55					60	
Leu	Val	Lys	Gly	Gly	Ile	Ala	Ser	Asp	Leu	Ile	Glu	Ile	Ile	Gly	
			65						70					75	
Thr	Ser	Tyr	Tyr	His	Ala	Ile	Leu	Pro	Leu	Leu	Pro	Leu	Ser	Arg	
			80						85					90	
Val	Glu	Ala	Glu	Val	Gln	Arg	Asp	Arg	Glu	Val	Lys	Glu	Glu	Leu	
			95						100					105	
Phe	Glu	Val	Ser	Pro	Lys	Gly	Phe	Trp	Leu	Pro	Glu	Leu	Ala	Tyr	
			110						115					120	
Asn	Pro	Ile	Ile	Phe	Ala	Ile	Leu	Lys	Asp	Asn	Gly	Tyr	Glu	Tyr	
			125						130					135	
Leu	Phe	Ala	Asp	Gly	Glu	Ala	Met	Leu	Phe	Ser	Ala	His	Leu	Asn	
			140						145					150	

005720-23057959

Ser	Ala	Ile	Lys	Pro	Ile	Lys	Pro	Leu	Tyr	Pro	His	Leu	Ile	Lys	155	160	165
Ala	Gln	Arg	Glu	Lys	Arg	Phe	Arg	Tyr	Ile	Ser	Tyr	Leu	Leu	Gly	170	175	180
Leu	Arg	Glu	Leu	Arg	Lys	Ala	Ile	Lys	Leu	Val	Phe	Glu	Gly	Lys	185	190	195
Val	Thr	Leu	Lys	Ala	Val	Lys	Asp	Ile	Glu	Ala	Val	Pro	Val	Trp	200	205	210
Val	Ala	Val	Asn	Thr	Ala	Val	Met	Leu	Gly	Ile	Gly	Arg	Leu	Pro	215	220	225
Leu	Met	As?	Pro	Lys	Lys	Val	Ala	Ser	Trp	Ile	Glu	Asp	Lys	Asp	230	235	240
Asn	Ile	Leu	Leu	Tyr	Gly	Thr	Asp	Ile	Glu	Phe	Ile	Gly	Tyr	Arg	245	250	255
Asp	Ile	Ala	Gly	Tyr	Arg	Met	Ser	Val	Glu	Gly	Leu	Leu	Glu	Val	260	265	270
Ile	Asp	Glu	Leu	Asn	Ser	Glu	Leu	Cys	Leu	Pro	Ser	Glu	Leu	Lys	275	280	285
His	Ser	Gly	Arg	Glu	Leu	Tyr	Leu	Arg	Thr	Ser	Ser	Trp	Ala	Pro	290	295	300
Asp	Lys	Ser	Leu	Arg	Ile	Trp	Arg	Glu	Asp	Glu	Gly	Asn	Ala	Arg	305	310	315
Leu	Asn	Met	Leu	Ser	Tyr	Asn	Met	Arg	Gly	Glu	Leu	Ala	Phe	Leu	320	325	330
Ala	Glu	Asn	Ser	Asp	Ala	Arg	Gly	Trp	Glu	Pro	Leu	Pro	Gln	Arg	335	340	345
Arg	Leu	Asp	Ala	Phe	Arg	Ala	Ile	Tyr	Asn	Asp	Trp	Arg	Gly	Glu	350	355	360
Asn	Gly	Glu	Pro														

WE CLAIM:

1. A method for hydrolyzing  $\alpha$ -galactose bonds comprising:  
contacting a compound having an  $\alpha$ -galactose bond with an effective amount of an enzyme having at least a 70% amino acid identity to amino acid sequence set forth in SEQ ID NO: 4.
2. The method according to claim 1 wherein the enzyme has at least 90% amino acid identity to the amino acid sequence set forth in SEQ ID NO: 4.
3. The method according to claim 1 wherein the enzyme comprises a sequence of amino acids identical to amino acids 1 to 364 of SEQ ID NO:4.
4. The method according to claim 1 wherein the enzyme has the amino acid sequence as set forth in SEQ ID NO: 4.
5. The method according to claim 1 wherein the enzyme is recombinantly produced.
6. The method according to claim 1 wherein the compound having the  $\alpha$ -galactose bond is raffinose.
7. The method according to claim 6 wherein the  $\alpha$ -galactose bond is in raw beet sugar.
8. The method according to claim 1 wherein the compound is raffinose, stachyose, verbascose, or a combination thereof.
9. The method according to claim 8 wherein the compound is contained in a member of the lentil or bean family, or a combination thereof.

10. The method according to claim 1 wherein the contacting is at a temperature of about 85° C.

11. The method according to claim 1 wherein the contacting is at a pH of about 9.5.

12. The method according to claim 1 wherein the contacting is at a temperature of about 85° C and a pH of about 9.5.

### ABSTRACT

A thermostable alpha-glycosidase derived from various *Thermococcus*, *alcaliphilus AEDII12RA* is disclosed. The enzymes are produced from native or recombinant host cells and can be utilized in the food processing industry.

096740-260960



FIG. 1

TTG Leu	AGA Arg	GCG Ala	CTC Leu	GTC Val <sup>5</sup>	TTT Phe	CAC His	GGC Gly	AAC Asn	CTC Leu <sup>10</sup>	CAG Gln	TAT Tyr	GCC Ala	GAA Glu	ATC Ile <sup>15</sup>	CCA Pro	48
AAG Lys	AGC Ser	GAA Glu	ATC Ile <sup>20</sup>	CCA Pro	AAG Lys	GTC Val	ATA Ile	GAG Glu <sup>25</sup>	AAG Lys	GCA Ala	TAC Tyr	ATC Ile	CCA Pro <sup>30</sup>	GTC Val	ATC Ile	96
GAG Glu	ACA Thr	CTG Leu <sup>35</sup>	ATT Ile	AAA Lys	GAA Glu	GAA Glu	ATT Ile <sup>40</sup>	CCT Pro	TTT Phe	GGG Gly	CTC Leu	AAC Asn <sup>45</sup>	ATA Ile	ACG Thr	GGC Gly	144
TAT Tyr	ACC Thr <sup>50</sup>	TTA Leu	AAG Lys	TTC Phe	CTC Leu	CCG Pro <sup>55</sup>	AAG Lys	GAT Asp <sup>60</sup>	ATT Ile	ATA Ile	GAC Asp	CTC Leu	GTT Val	AAA Lys	GGG Gly	192
GGC Gly <sup>65</sup>	ATC Ile	GCG Ala	AGT Ser	GAC Asp	CTG Leu <sup>70</sup>	ATA Ile	GAG Glu	ATA Ile	ATC Ile	GGA Gly <sup>75</sup>	ACG Thr	AGC Ser	TAC Tyr	ACG Thr	CAC His <sup>80</sup>	240
GCA Ala	ATA Ile	CTC Leu	CCC Pro	CTC Leu <sup>85</sup>	CTC Leu	CCG Pro	CTT Leu	AGC Ser	AGA Arg <sup>90</sup>	GTA Val	GAA Glu	GCA Ala	CAA Gln	GTT Val <sup>95</sup>	CAG Gln	288
AGA Arg	GAT Asp	AGG Arg	GAA Glu <sup>100</sup>	GTT Val	AAG Lys	GAA Glu	GAG Glu	CTC Leu <sup>105</sup>	TTC Phe	GAG Glu	CTT Val	TCT Ser	CCA Pro <sup>110</sup>	AAG Lys	GGA Gly	336
TTC Phe	TGG Trp	CTG Leu <sup>115</sup>	CCA Pro	GAG Glu	CTC Leu	GCC Ala	TAT Tyr <sup>120</sup>	GAC Asp	CCG Pro	ATA Ile	ATC Ile	CCT Pro <sup>125</sup>	GCC Ala	ATA Ile	CTG Leu	384

FIG. 1A

AAG GAC AAC GGT TAT GAG TAT CTA TTC GCC GAC GGG GAG GCG ATG CTT	432
Lys Asp 130 Asn Gly Tyr 135 Glu Tyr Leu Phe Ala Asp Gly 140	
TTC TCA GCT CAT CTC AAC AAC TCG GCG ATA AAG CCA ATT AAA CCG CTC TAT	480
Phe Ser Ala His Leu Asn 150 Ser Ala Ile Lys Pro Ile Lys Pro Leu Tyr 160	
CCA CAC CTT ATA AAG GCC CAA AGG GAA AAG AAG CGC TTT AGG TAC ATC AGC	528
Pro His 3Leu Ile Lys 165 Ala Gln Arg Glu Lys Lys Arg Phe Ile Tyr 175 Ser	
TAT CTC CTT GGT CTC AGG GAG GAG CTT AGG AAG AAG GCG ATA AAG AAG CTC GTT	576
Tyr Leu Leu Gly 180 Arg Glu Leu Arg Lys Lys Ala Ile Lys Lys Leu Val Phe	
GAA GGT AAG GTA ACG CTA AAG AAG GTC GTC AAA GAC ATC GAA GCC GTA CCC	624
Glu Gly Lys 195 Thr Leu Lys Lys Val Lys Asp Ile Glu 205 Ala Val Pro	
GTT TGG GTG GCC GTG AAC ACG GCT GTA ATG CTC GGC ATC GGA AGG CTT	672
Val Trp 210 Val Ala Val Asn Thr 215 Ala Val Met Leu Gly Ile Ile Gly Arg Leu	
CCT CTT ATG AAT CCT AAG AAA GTG GCG AGC TGG ATA GAG GAC AAG GAC	720
Pro Leu Met Asn Pro Lys 230 Lys Lys Ala Ser Trp Ile Glu Asp Lys 240	
AAC ATT CTT CTA TAC GGC ACC GAT ATA GAG TTC ATT GGC TAT AGG GAC	768
Asn Ile Leu Leu Tyr 245 Gly Thr Asp Ile Ile Glu Phe Ile Gly Tyr Arg 255 Asp	

FIG. 1B

ATT GCA GGC TAC AGA ATG AGT GTT GAG GGA TTA TTA GAG GAG GAT ATA GAC	816
Ile Ala Gly Tyr Arg Met Ser Val Glu Gly Leu Leu Glu Val Ile Asp	
GAG CTC AAC TCG GAA CTG TGC CTT CCC TCA GAG CTG AAG CAC ATG GGA	864
Glu Leu Asn Ser Glu Glu Leu Cys Leu Pro Ser Glu Leu Lys His Ser Gly	
AGG GAG CTC TAC TAC TTA CGG ACT TCG AGT TGG GCA CCA GAT AAG AGC TTG	912
Arg Glu Leu Tyr Tyr Leu Arg Thr Ser Ser Trp Ala Pro Asp Lys Ser Leu	
AGG ATA TGG AGA GAG GAC GAA GGG AAC GCA AGA CTT AAT ATG CTG TCC	960
Arg Ile Trp Arg Glu Glu Asp Glu Gly Asn Ala Arg Leu Asn Met Leu Ser	
TAC AAT ATG AGG GGC GAA CTC GCC CTT TTA GCC GAG AAC AGC GAT GCA	1008
Tyr Asn Met Arg Gly Glu Leu Leu Leu Phe Leu Ala Glu Asn Ser Asp Ala	
AGG GGA TGG GAG CCC CTC CCT GAG AGG AGG CTG GAT GCC TTC CGG GCG	1047
Arg Gly Trp Glu Pro Leu Pro Glu Arg Arg Leu Asp Ala Phe Arg Ala	
ATA TAT AAC GAT TGG AGG GGT GAA AAT GGG GAA CCT TAG	1086
Ile Tyr Asn Asp Trp Arg Gly Glu Glu Pro Pro End	
	355
	360
	365
	370
	375
	380
	385
	390
	395
	400
	405
	410
	415
	420
	425
	430
	435
	440
	445
	450
	455
	460
	465
	470
	475
	480
	485
	490
	495
	500
	505
	510
	515
	520
	525
	530
	535
	540
	545
	550
	555
	560
	565
	570
	575
	580
	585
	590
	595
	600
	605
	610
	615
	620
	625
	630
	635
	640
	645
	650
	655
	660
	665
	670
	675
	680
	685
	690
	695
	700
	705
	710
	715
	720
	725
	730
	735
	740
	745
	750
	755
	760
	765
	770
	775
	780
	785
	790
	795
	800
	805
	810
	815
	820
	825
	830
	835
	840
	845
	850
	855
	860
	865
	870
	875
	880
	885
	890
	895
	900
	905
	910
	915
	920
	925
	930
	935
	940
	945
	950
	955
	960
	965
	970
	975
	980
	985
	990
	995
	1000
	1005
	1010
	1015
	1020
	1025
	1030
	1035
	1040
	1045
	1050
	1055
	1060
	1065
	1070
	1075
	1080
	1085
	1090
	1095
	1100
	1105
	1110
	1115
	1120
	1125
	1130
	1135
	1140
	1145
	1150
	1155
	1160
	1165
	1170
	1175
	1180
	1185
	1190
	1195
	1200
	1205
	1210
	1215
	1220
	1225
	1230
	1235
	1240
	1245
	1250
	1255
	1260
	1265
	1270
	1275
	1280
	1285
	1290
	1295
	1300
	1305
	1310
	1315
	1320
	1325
	1330
	1335
	1340
	1345
	1350
	1355
	1360
	1365
	1370
	1375
	1380
	1385
	1390
	1395
	1400
	1405
	1410
	1415
	1420
	1425
	1430
	1435
	1440
	1445
	1450
	1455
	1460
	1465
	1470
	1475
	1480
	1485
	1490
	1495
	1500
	1505
	1510
	1515
	1520
	1525
	1530
	1535
	1540
	1545
	1550
	1555
	1560
	1565
	1570
	1575
	1580
	1585
	1590
	1595
	1600
	1605
	1610
	1615
	1620
	1625
	1630
	1635
	1640
	1645
	1650
	1655
	1660
	1665
	1670
	1675
	1680
	1685
	1690
	1695
	1700
	1705
	1710
	1715
	1720
	1725
	1730
	1735
	1740
	1745
	1750
	1755
	1760
	1765
	1770
	1775
	1780
	1785
	1790
	1795
	1800
	1805
	1810
	1815
	1820
	1825
	1830
	1835
	1840
	1845
	1850
	1855
	1860
	1865
	1870
	1875
	1880
	1885
	1890
	1895
	1900
	1905
	1910
	1915
	1920
	1925
	1930
	1935
	1940
	1945
	1950
	1955
	1960
	1965
	1970
	1975
	1980
	1985
	1990
	1995
	2000
	2005
	2010
	2015
	2020
	2025
	2030
	2035
	2040
	2045
	2050
	2055
	2060
	2065
	2070
	2075
	2080
	2085
	2090
	2095
	2100
	2105
	2110
	2115
	2120
	2125
	2130
	2135
	2140
	2145
	2150
	2155
	2160
	2165
	2170
	2175
	2180
	2185
	2190
	2195
	2200
	2205
	2210
	2215
	2220
	2225
	2230
	2235
	2240
	2245
	2250
	2255
	2260
	2265
	2270
	2275
	2280
	2285
	2290
	2295
	2300
	2305
	2310
	2315
	2320
	2325
	2330
	2335
	2340
	2345
	2350
	2355
	2360
	2365
	2370
	2375
	2380
	2385
	2390
	2395
	2400
	2405
	2410
	2415
	2420
	2425
	2430
	2435
	2440
	2445
	2450
	2455
	2460
	2465
	2470
	2475
	2480
	2485
	2490
	2495
	2500
	2505
	2510
	2515
	2520
	2525
	2530
	2535
	2540
	2545
	2550
	2555
	2560
	2565
	2570
	2575
	2580
	2585
	2590
	2595
	2600
	2605
	2610
	2615
	2620
	2625
	2630
	2635
	2640
	2645
	2650
	2655
	2660
	2665
	2670
	2675
	2680
	2685
	2690
	2695
	2700
	2705
	2710
	2715
	2720
	2725
	2730
	2735
	2740
	2745
	2750
	2755
	2760
	2765
	2770
	2775
	2780
	2785
	2790
	2795
	2800
	2805
	2810
	2815
	2820
	2825
	2830
	2835
	2840
	2845
	2850
	2855
	2860
	2865
	2870
	2875
	2880
	2885
	2890
	2895
	2900
	2905
	2910
	2915
	2920
	2925
	2930
	2935
	2940
	2945
	2950
	2955
	2960
	2965
	2970
	2975
	2980
	2985
	2990
	2995
	3000
	3005
	3010
	3015
	3020
	3025
	3030
	3035
	3040
	3045
	3050
	3055
	3060
	3065
	3070
	3075
	3080
	3085
	3090
	3095
	3100
	3105
	3110
	3115
	3120
	3125
	3130
	3135
	3140
	3145
	3150
	3155
	3160
	3165
	3170
	3175
	3180
	3185
	3190
	3195
	3200
	3205
	3210
	3215
	3220
	3225
	3230
	3235
	3240
	3245
	3250
	3255
	3260
	3265
	3270
	3275
	3280
	3285
	3290
	3295
	3300
	3305
	3310
	3315
	3320
	3325
	3330
	3335
	3340
	3345
	3350
	3355
	3360
	3365
	3370
	3375
	3380
	3385
	3390
	3395
	3400
	3405
	3410
	3415
	3420
	3425
	3430
	3435
	3440

DECLARATION FOR PATENT APPLICATION

As a below-named inventor, I hereby declare that:

My residence, post office address and citizenship is as stated below next to my name.

I believe that I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled ALPHA-GALACTOSIDASE, the specification of which

\_\_\_\_\_ is attached hereto.

X was filed on July 19, 2000 (Attorney Docket No. DIVER1120-3)  
as U.S. Application Serial No. \_\_\_\_\_  
and was amended on \_\_\_\_\_  
if applicable (the "Application").

I hereby authorize and request insertion of the application serial number of the Application when officially known.

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability of the subject matter of the Application as defined in Title 37, Code of Federal Regulations ("C.F.R."), § 1.56.

With respect to the Application, I hereby claim the benefit under 35 U.S.C. Section 120 of any United States application(s), or Section 365(c) of any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of the application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of 35 U.S.C. Section 112, I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability of the subject matter of the Application as defined in Title 37, C.F.R., Section 1.56 which became available between the filing date of the prior application and the national or PCT International filing date of the Application:

<u>09/407,806</u> (Application Serial No.)	<u>September 28, 1999</u> (Filing Date)	<u>Pending</u> (Status)
<u>08/613,220</u> (Application Serial No.)	<u>March 8, 1996</u> (Filing Date)	<u>Issued September 28, 1999 as</u> <u>U.S. Patent No. 5,958,751</u>

I hereby claim foreign priority benefits under Title 35, United States Code, § 119 of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) of which priority is claimed:

COUNTRY	APPLICATION NO.	FILING DATE	PRIORITY CLAIMED
_____	_____	_____	<input type="checkbox"/> Yes <input type="checkbox"/> No

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of first inventor: **Dennis Murphy**

Inventor's signature: \_\_\_\_\_

Date: \_\_\_\_\_

Residence: Paoli, PA 19301

Citizenship: United States

Post Office Address: 10 Fairway Road  
Paoli, PA 19301

Full name of second inventor: **John Reid**

Inventor's signature: \_\_\_\_\_

Date: \_\_\_\_\_

Residence: Bryn Mawr, PA 19010

Citizenship: United States

Post Office Address: 922 Montgomery Ave., Apt. J-2  
Bryn Mawr, PA 19010

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	Murphy and Reid	Art Unit:	Unassigned
Serial No.:		Examiner	Unassigned
Filed:	July 19, 2000		
Title:	ALPHA-GALACTOSIDASE		

Assistant Commissioner for Patents  
Washington, D.C. 20231

**REVOCATION OF POWERS OF ATTORNEY;**  
**POWER OF ATTORNEY BY ASSIGNEE AND CHANGE OF ADDRESS**

Sir:

**DIVERSA CORPORATION**, a Delaware corporation, the assignee of the entire right, title and interest in the above-identified application hereby revokes all previous Powers of Attorney and appoints the following attorneys to prosecute the above-identified patent application and to transact all business in the Patent and Trademark Office connected therewith:

In re Application of Murphy and Reid  
Application Serial No.:  
Filed: July 19, 2000  
Page 2

PATENT  
Attorney Docket No.: DIVER1120-3

JOSEPH R. BAKER	Registration No. 40,900
TIM ELLIS	Registration No. 41,734
LISA A. HAILE	Registration No. 38,347
WILLIAM N. HULSEY, III	Registration No. 33,402
RICHARD J. IMBRA	Registration No. 37,643
SHEILA R. KIRSCHENBAUM	Registration No. 44,835
JUNE M. LEARN	Registration No. 31,238
TIMOTHY W. LOHSE	Registration No. 35,255
TERRANCE A. MEADOR	Registration No. 30,298
JOHN OSKOREP	Registration No. 41,234
STEPHEN E. REITER	Registration No. 31,192
STEVEN R. SPRINKLE	Registration No. 40,825
RAMSEY R. STEWART	Registration No. 38,322
DAVID R. STEVENS	Registration No. 38,626
BARRY N. YOUNG	Registration No. 27,744

We hereby state that we have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment(s) referred to above.

We authorize and request insertion of the serial number of this application when officially known.

We acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, Section 1.56(a).



In re Application of Murphy and Reid  
Application Serial No.:  
Filed: July 19, 2000  
Page 3

PATENT  
Attorney Docket No.: DIVER1120-3

Direct all telephone calls and correspondence to:

LISA A. HAILE, Ph.D.  
GRAY CARY WARE & FREIDENRICH LLP  
4365 Executive Drive, Suite 1600  
San Diego, CA 92121  
Direct Line: (858) 677-1456  
Facsimile: (858) 677-1465

**DIVERSA CORPORATION**  
a Delaware corporation

By: \_\_\_\_\_

Name: Carolyn Erickson

Title: Vice President, Intellectual Property

Date: \_\_\_\_\_

DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

ALPHA-GALACTOSIDASE

the specification of which [X] is attached hereto or [ ] was filed on \_\_\_\_\_ as Application Serial No. \_\_\_\_\_ and was amended on \_\_\_\_\_ (if applicable).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the patentability of this application in accordance with Title 37, Code of Federal Regulations, Section 1.56(a).

I hereby claim foreign priority benefits under Title 35, United States Code, Section 119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed. Prior Foreign Application(s):

Priority Claimed

(Number)	(Country)	(Day/Month/Year Filed)	Yes	No
			<input type="checkbox"/>	<input type="checkbox"/>

I hereby claim the benefit under Title 35, United States Code, Section 120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, Section 112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, Section 1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

(Application Serial No.)	(Filing Date)	Pending (Status - patented, pending, abandoned)

I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith: John N. Bain (Reg. No. 18,651); John G. Giffillan, III (Reg. No. 22,746); Elliot M. Olstein (Reg. No. 24,025); Raymond J. Lillie (Reg. No. 31,778); Charles J. Herron (Reg. No. 28,019); William Squire (Reg. No. 25,378); Kenneth S. Weitzman (Reg. No. 36,306); and Gregory Ferraro (Reg. No. 36,134). Address correspondence and telephone calls to Charles J. Herron c/o Carella, Byrne, Bain, Giffillan, Cecchi, Stewart & Olstein, 6 Becker Farm Road, Roseland, NJ 07068 - (201) 994-1700.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

Full name of inventor: Dennis Murphy  
 Inventor's signature: *Dennis Murphy* Date: Mar. 5, 1996  
 Residence: 10 Fairway Road, Paoli, PA 19301 Citizenship: United States  
 Post Office Address: same

Full name of inventor: John Reid  
 Inventor's signature: *John C. Reid* Date: Mar. 5, 1996  
 Residence: 922 Montgomery Avenue Apt. J-2, Bryn Mawr, PA 19010 Citizenship: United States  
 Post Office Address: same

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants: Murphy, D. *et al.* Art Unit: 1814  
Serial No.: 08/613,220 Examiner: Hendricks, K.  
Filed: 3/8/96  
Title: ALPHA-GALACTOSIDASE

Assistant Commissioner for Patents  
Washington, DC 20231

REVOCATION AND NEW POWER OF ATTORNEY

Under 37 CFR §3.73(b) RECOMBINANT BIOCATALYSIS, INC., a Delaware corporation, certifies that it is the assignee of 100% of the right, title and interest in the patent application identified above by virtue of an assignment from the inventors of the patent application identified above. The assignment was recorded in the Patent and Trademark Office at Reel 8103, Frames 0390-0393, on July 18, 1996.

The undersigned has reviewed all the documents in the chain of title of the patent application identified above and, to the best of undersigned's knowledge and belief, title is in the assignee identified above.

The undersigned, whose title is supplied below, is empowered to act on behalf of the assignee.

Date of Deposit 6-13-97  
I hereby certify under 37 CFR 1.8(a) that this correspondence is being deposited with the United States Postal Service as first class mail with sufficient postage on the date indicated above and is addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231.

Stephanie Sharrett  
Stephanie Sharrett

096130220-071900

The undersigned, acting on behalf of the assignee, hereby revokes all powers of attorney previously granted in the application and appoints: John R. Wetherell, Jr., Ph.D. (Reg. No. 31,678); Lisa A. Haile (Reg. No. 38,347); Stacy L. Taylor (Reg. No. 34,842); John Land (Reg. No. 29,554), and June M. Learn (Reg. No. 31,238), of the firm of FISH & RICHARDSON P.C., as its attorneys with full power of substitution and revocation, to prosecute the application and to transact all business in the United States Patent and Trademark Office connected therewith.

Please direct all telephone calls to John R. Wetherell at (619) 678-5070 and all correspondence relative to said application to the following address:

John R. Wetherell, Jr., Ph.D.  
FISH & RICHARDSON, P.C.  
4225 Executive Square, Suite 1400  
La Jolla, California 92037  
(619) 678-5099 (Facsimile)

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patents issued thereon.

Respectfully submitted,

Date: 5-17-97



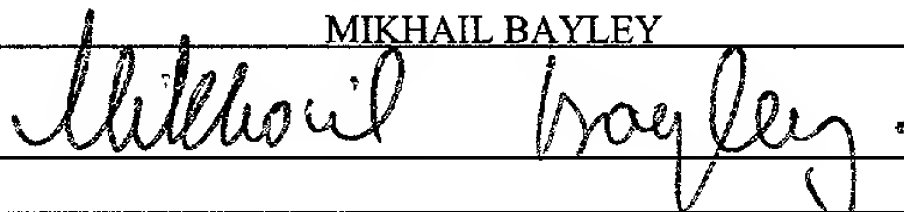
Carolyn Erickson  
Manager, Business Development and  
Regulatory Affairs  
RECOMBINANT BIOCATALYSIS, INC.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Murphy and Reid  
Parent Serial No. 09/407,806  
Parent Filing Date September 28, 1999  
Serial No.:  
Filed: July 19, 2000  
Title: ALPHA-GALACTOSIDASE  
Art Unit: Unassigned  
Examiner: Unassigned

**Box Patent Application**

Assistant Commissioner for Patents  
Washington, D.C. 20231

"EXPRESS MAIL" Mailing Label Number: **EL476991301US**  
Deposit **July 19, 2000**  
I hereby certify under 37 CFR 1.10 that this correspondence is being deposited with the United States Postal Service as "Express Mail Post Office to Addressee" with sufficient postage on the date indicated above and is addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231  
  
MIKHAIL BAYLEY

**PERMISSION TO USE SEQUENCE LISTING**

Sir:

The above-identified patent application lacks a substitute paper copy of the Sequence Listing for inclusion into the Specification, as well as a computer readable form of the Sequence Listing. Applicants respectfully direct the attention of the Office to the following:

1. A complete paper copy of the Sequence Listing is to be inserted following the Abstract on page 27 and before the Drawings. Please insert the Sequence Listing beginning with page 1 and numbering consecutively thereafter. A paper copy of the Sequence Listing is included herewith, and is identical to the computer readable copy of the Sequence Listing filed in U.S. Patent Application Number 09/407,806 (the "806 application"), filed September 28, 1999.


2. A computer-readable form in this application is identical with that filed in '806 application, filed September 28, 1999. Pursuant to 37 CFR §1.821(e), please use the last-filed computer-readable form filed in the '806 application as the computer-readable form for application filed herewith. It is understood that the Patent and Trademark Office will make the necessary change in application number and filing date for the computer readable form that will be used for application filed herewith.

I hereby state, as required by 37 C.F.R. § 1.821(g), that the enclosed submission includes no new matter. Applicants submit that the foregoing satisfies the requirements of Rule §1.821. If there are any questions regarding this response, the Office is invited to contact the undersigned.

No fee is deemed necessary in connection with the filing of this paper. However, if any fee is required, the Commissioner is hereby authorized to charge the amount of this fee, or credit any overpayments, to Deposit Account No. 50-1355.

Respectfully submitted,

Date: 7/19/00

  
Lisa A. Haile, Ph.D.  
Registration No. 38,347  
Telephone: (858) 677-1456  
Facsimile: (858) 677-1465

GRAY CARY WARE & FREIDENRICH LLP  
4365 Executive Drive, Suite 1600  
San Diego, California 92121-2189

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: MURPHY, Dennis  
REID, John
- (ii) TITLE OF INVENTION: Alpha Glycosidase
- (iii) NUMBER OF SEQUENCES: 4
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
CECCHI, STEWART & OLSTEIN
- (B) STREET: 6 BECKER FARM ROAD
- (C) CITY: ROSELAND
- (D) STATE: NEW JERSEY
- (E) COUNTRY: USA
- (F) ZIP: 07068
- (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: 3.5 INCH DISKETTE
- (B) COMPUTER: IBM PS/2
- (C) OPERATING SYSTEM: MS-DOS
- (D) SOFTWARE: ASCII
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER: Unassigned
- (B) FILING DATE: Concurrently
- (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER:
- (B) FILING DATE:
- (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: HERRON, CHARLES J.
- (B) REGISTRATION NUMBER: 28,019
- (C) REFERENCE/DOCKET NUMBER: 331400-40
- (ix) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: 201-894-1700
- (B) TELEFAX: 201-994-1744
- (2) INFORMATION FOR SEQ ID NO:1:
- (i) SEQUENCE CHARACTERISTICS
- (A) LENGTH: 52 NUCLEOTIDES
- (B) TYPE: NUCLEIC ACID
- (C) STRANDEDNESS: SINGLE
- (D) TOPOLOGY: LINEAR
- (ii) MOLECULE TYPE: cDNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
- CCGAGAATTC ATTAAAGAGG AGAAATTAAC TATGAGAGCG CTCGTCTTTC AC
- (2) INFORMATION FOR SEQ ID NO:2:

52

(i) SEQUENCE CHARACTERISTICS  
 (A) LENGTH: 31 NUCLEOTIDES  
 (B) TYPE: NUCLEIC ACID  
 (C) STRANDEDNESS: SINGLE  
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CGGAAGATCT AGGTTCCCA TTTTCACCCC T

31

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS  
 (A) LENGTH: 1,095 NUCLEOTIDES  
 (B) TYPE: NUCLEIC ACID  
 (C) STRANDEDNESS: SINGLE  
 (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TTG AGA GCG CTC GTC TTT CAC GGC AAC CTC CAG TAT GCC GAA ATC CCA 48  
 Leu Arg Ala Leu Val Phe His Gly Asn Leu Gln Tyr Ala Glu Ile Pro  
                   5                  10                  15

AAG AGC GAA ATC CCA AAG GTC ATA GAG AAG GCA TAC ATC CCA GTC ATC 96  
 Lys Ser Glu Ile Pro Lys Val Ile Glu Lys Ala Tyr Ile Pro Val Ile  
                   20                  25                  30

GAG ACA CTG ATT AAA GAA GAA ATT CCT TTT GGG CTC AAC ATA ACG GGC 144  
 Glu Thr Leu Ile Lys Glu Glu Ile Pro Phe Gly Leu Asn Ile Thr Gly  
                   35                  40                  45

TAT ACC TTA AAG TTC CTC CCG AAG GAT ATT ATA GAC CTC GTT AAA GGG 192  
 Tyr Thr Leu Lys Phe Leu Pro Lys Asp Ile Ile Asp Leu Val Lys Gly  
                   50                  55                  60

GGC ATC GCG AGT GAC CTG ATA GAG ATA ATC GGA ACG AGC TAC ACG CAC 240  
 Gly Ile Ala Ser Asp Leu Ile Glu Ile Ile Gly Thr Ser Tyr Thr His  
                   65                  70                  75                  80

GCA ATA CTC CCC CTC CTC CCG CTT AGC AGA GTA GAA GCA CAA GTT CAG 288  
 Ala Ile Leu Pro Leu Leu Pro Leu Ser Arg Val Glu Ala Gln Val Gln  
                   85                  90                  95

AGA GAT AGG GAA GTT AAG GAA GAG CTC TTC GAG CTT TCT CCA AAG GGA 336  
 Arg Asp Arg Glu Val Lys Glu Glu Leu Phe Glu Val Ser Pro Lys Gly  
                   100                  105                  110



TTC	TGG	CTG	CCA	GAG	CTC	GCC	TAT	GAC	CCG	ATA	ATC	CCT	GCC	ATA	CTG	384
Phe	Trp	Leu	Pro	Glu	Leu	Ala	Tyr	Asp	Pro	Ile	Ile	Pro	Ala	Ile	Leu	
		115					120					125				
AAG	GAC	AAC	GGT	TAT	GAG	TAT	CTA	TTC	GCC	GAC	GGG	GAG	GCG	ATG	CTT	432
Lys	Asp	Asn	Gly	Tyr	Glu	Tyr	Leu	Phe	Ala	Asp	Gly	Glu	Ala	Met	Leu	
	130					135					140					
TTC	TCA	GCT	CAT	CTC	AAC	TCG	GCG	ATA	AAG	CCA	ATT	AAA	CCG	CTC	TAT	480
Phe	Ser	Ala	His	Leu	Asn	Ser	Ala	Ile	Lys	Pro	Ile	Lys	Pro	Leu	Tyr	
145					150					155					160	
CCA	CAC	CTT	ATA	AAG	GCC	CAA	AGG	GAA	AAG	CGC	TTT	AGG	TAC	ATC	AGC	528
Pro	His	3Leu	Ile	Lys	Ala	Gln	Arg	Glu	Lys	Arg	Phe	Arg	Tyr	Ile	Ser	
				165				170					175			
TAT	CTC	CTT	GGT	CTC	AGG	GAG	CTT	AGG	AAG	GCG	ATA	AAG	CTC	GTT	TTT	576
Tyr	Leu	Leu	Gly	Leu	Arg	Glu	Leu	Arg	Lys	Ala	Ile	Lys	Leu	Val	Phe	
			180					185					190			
GAA	GGT	AAG	GTA	ACG	CTA	AAG	GCA	GTC	AAA	GAC	ATC	GAA	GCC	GTA	CCC	624
Glu	Gly	Lys	Val	Thr	Leu	Lys	Ala	Val	Lys	Asp	Ile	Glu	Ala	Val	Pro	
		195					200					205				
GTT	TGG	GTG	GCC	GTG	AAC	ACG	GCT	GTA	ATG	CTC	GGC	ATC	GGA	AGG	CTT	672
Val	Trp	Val	Ala	Val	Asn	Thr	Ala	Val	Met	Leu	Gly	Ile	Gly	Arg	Leu	
	210					215					220					
CCT	CTT	ATG	AAT	CCT	AAG	AAA	GTG	GCG	AGC	TGG	ATA	GAG	GAC	AAG	GAC	720
Pro	Leu	Met	Asn	Pro	Lys	Lys	Val	Ala	Ser	Trp	Ile	Glu	Asp	Lys	Asp	
225					230					235					240	
AAC	ATT	CTT	CTA	TAC	GGC	ACC	GAT	ATA	GAG	TTC	ATT	GGC	TAT	AGG	GAC	768
Asn	Ile	Leu	Leu	Tyr	Gly	Thr	Asp	Ile	Glu	Phe	Ile	Gly	Tyr	Arg	Asp	
				245				250						255		
ATT	GCA	GGC	TAC	AGA	ATG	AGT	GTT	GAG	GGA	TTA	TTA	GAG	GTT	ATA	GAC	816
Ile	Ala	Gly	Tyr	Arg	Met	Ser	Val	Glu	Gly	Leu	Leu	Glu	Val	Ile	Asp	
			260					265					270			
GAG	CTC	AAC	TCG	GAA	CTG	TGC	CTT	CCC	TCA	GAG	CTG	AAG	CAC	AGT	GGA	864
Glu	Leu	Asn	Ser	Glu	Leu	Cys	Leu	Pro	Ser	Glu	Leu	Lys	His	Ser	Gly	
		275					280					285				
AGG	GAG	CTC	TAC	TTA	CGG	ACT	TCG	AGT	TGG	GCA	CCA	GAT	AAG	AGC	TTG	912
Arg	Glu	Leu	Tyr	Leu	Arg	Thr	Ser	Ser	Trp	Ala	Pro	Asp	Lys	Ser	Leu	
	290					295					300					

AGG	ATA	TGG	AGA	GAG		GAA	GGG	AAC	GCA	AGA	CTT	AAT	ATC	CTG	TCC	960
Arg	Ile	Trp	Arg	Glu		Glu	Gly	Asn	Ala	Arg	Leu	Asn	M	Leu	Ser	
305					310					315					320	
TAC	AAT	ATG	AGG	GGC	GAA	CTC	GCC	CTT	TTA	GCC	GAG	AAC	AGC	GAT	GCA	1008
Tyr	Asn	Met	Arg	Gly	Glu	Leu	Ala	Phe	Leu	Ala	Glu	Asn	Ser	Asp	Ala	
				325					330					335		
AGG	GGA	TGG	GAG	CCC	CTC	CCT	GAG	AGG	AGG	CTG	GAT	GCC	TTC	CGG	GCG	1047
Arg	Gly	Trp	Glu	Pro	Leu	Pro	Glu	Arg	Arg	Leu	Asp	Ala	Phe	Arg	Ala	
			340					345					350			
ATA	TAT	AAC	GAT	TGG	AGG	GGT	GAA	AAT	GGG	GAA	CCT	TAG				1086
Ile	Tyr	Asn	Asp	Trp	Arg	Gly	Glu	Asn	Gly	Glu	Pro	End				
		355					360					365				

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 364 AMINO ACIDS
- (B) TYPE: POLYPEPTIDE
- (D) TOPOLOGY: LINEAR

(ii) MOLECULE TYPE: PROTEIN

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Leu	Arg	Ala	Leu	Val	Phe	Mis	Gly	Asn	Leu	Gln	Tyr	Ala	Glu	Ile	
			5					10						15	
Phe	Lys	Ser	Glu	Ile	Pro	Lys	Val	Ile	Glu	Lys	Ala	Tyr	Ile	Pro	
			20					25						30	
Val	Ile	Glu	Thr	Leu	Ile	Lys	Glu	Gln	Ile	Pro	Phe	Gly	Leu	Asn	
			35					40						45	
Ile	Thr	Gly	Tyr	Thr	Leu	Lys	Phe	Leu	Pro	Lys	Asn	Ile	Ile	Asp	
			50					55						60	
Leu	Val	Lys	Gly	Gly	Ile	Ala	Ser	Asp	Leu	Ile	Glu	Ile	Ile	Gly	
			65					70						75	
Thr	Ser	Tyr	Tyr	His	Ala	Ile	Leu	Pro	Leu	Leu	Pro	Leu	Ser	Arg	
			80					85						90	
Val	Glu	Ala	Glu	Val	Gln	Arg	Asp	Arg	Glu	Val	Lys	Glu	Glu	Leu	
			95					100						105	
Phe	Glu	Val	Ser	Pro	Lys	Gly	Phe	Trp	Leu	Pro	Glu	Leu	Ala	Tyr	
			110					115						120	
Asn	Pro	Ile	Ile	Phe	Ala	Ile	Leu	Lys	Asp	Asn	Gly	Tyr	Glu	Tyr	
			125					130						135	
Leu	Phe	Ala	Asp	Gly	Glu	Ala	Met	Leu	Phe	Ser	Ala	His	Leu	Asn	
			140					145						150	

005740" 2305T 950

Ser	Ala	Ile	Lys	Pro	Ile	Lys	Pro	Leu	Tyr	Pro	His	Leu	Ile	Lys	155	160	165
Ala	Gln	Arg	Glu	Lys	Arg	Phe	Arg	Tyr	Ile	Ser	Tyr	Leu	Leu	GLy	170	175	180
Leu	Arg	Glu	Leu	Arg	Lys	Ala	Ile	Lys	Leu	Val	Phe	Glu	Gly	Lys	185	190	195
Val	Thr	Leu	Lys	Ala	Val	Lys	Asp	Ile	Glu	Ala	Val	Pro	Val	Trp	200	205	210
Val	Ala	Val	Asn	Thr	Ala	Val	Met	Leu	Gly	Ile	Gly	Arg	Leu	Pro	215	220	225
Leu	Met	As?	Pro	Lys	Lys	Val	Ala	Ser	Trp	Ile	Glu	Asp	Lys	Asp	230	235	240
Asn	Ile	Leu	Leu	Tyr	Gly	Thr	Asp	Ile	Glu	Phe	Ile	Gly	Tyr	Arg	245	250	255
Asp	Ile	Ala	Gly	Tyr	Arg	Met	Ser	Val	Glu	Gly	Leu	Leu	Glu	Val	260	265	270
Ile	Asp	Glu	Leu	Asn	Ser	Glu	Leu	Cys	Leu	Pro	Ser	Glu	Leu	Lys	275	280	285
His	Ser	Gly	Arg	Glu	Leu	Tyr	Leu	Arg	Thr	Ser	Ser	Trp	Ala	Pro	290	295	300
Asp	Lys	Ser	Leu	Arg	Ile	Trp	Arg	Glu	Asp	Glu	Gly	Asn	Ala	Arg	305	310	315
Leu	Asn	Met	Leu	Ser	Tyr	Asn	Met	Arg	Gly	Glu	Leu	Ala	Phe	Leu	320	325	330
Ala	Glu	Asn	Ser	Asp	Ala	Arg	Gly	Trp	Glu	Pro	Leu	Pro	Gln	Arg	335	340	345
Arg	Leu	Asp	Ala	Phe	Arg	Ala	Ile	Tyr	Asn	Asp	Trp	Arg	Gly	Glu	350	355	360
Asn	Gly	Glu	Pro														